

# The accuracy of autobiographical memory: A replication of Barclay & Wellman (1986)

Citation for published version (APA):

Merckelbach, H. L. G. J., Wessel, I., & Horselenberg, R. (1997). The accuracy of autobiographical memory: A replication of Barclay & Wellman (1986). *Behavioural and Cognitive Psychotherapy*, 25, 103-111. <https://doi.org/10.1017/S1352465800018336>

## Document status and date:

Published: 01/01/1997

## DOI:

[10.1017/S1352465800018336](https://doi.org/10.1017/S1352465800018336)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# **The Accuracy of Autobiographical Memory: A Replication of Barclay & Wellman (1986)**

Harald Merckelbach, Ineke Wessel and  
Robert Horselenberg

Maastricht University, The Netherlands

Ten volunteers kept written records of self-selected, daily events for a one-week period. After four months, they were given a surprise recognition test. This test consisted of original memories and several types of foils. In line with previous findings of Barclay and Wellman, it was found that acceptance of foils as one's own memories is a relatively common phenomenon. This suggests that pseudomemories may also occur in the absence of suggestions and repeated interviews and, more generally, that autobiographical memory is intrinsically unreliable.

## **Introduction**

In the past three years or so, the issue of "recovered" (or "delayed") memories has stimulated a vast amount of debate and research (see, for reviews, Lindsay & Read, 1994; Loftus, 1993). In this context, several researchers and clinicians have argued that memories of childhood trauma recovered during treatment interventions such as hypnosis, guided imagery and the like may actually be pseudomemories (e.g. Loftus, 1993) and should be treated with "respectful scepticism" (Paris, 1995). Whether or not these authors are right, the debate on recovered memories has sensitized researchers and clinicians to the fallibility of autobiographical memory (see, for a review, Brewin, 1996). The relevance of this issue is not limited to treatment methods that focus on traumatized individuals, but extends to the broad context of psychotherapy. Indeed, as Trierweiler and Donovan (1994, p. 301) note in their thoughtful review: "memories of specific interpersonal events in a client's life are the grist for the psychotherapeutic mill".

So far, experimental studies concerned with memory reliability have emphasized the role that external factors play in memory illusions. There are two lines of research that illustrate this emphasis. In the first type of study, subjects are exposed to an event and afterwards receive misleading information about that event. Under this condition, some subjects come to believe information that was merely suggested to them. This so called

"post-event misinformation effect" (Loftus, 1979) appears to be a robust phenomenon that can be elicited in a wide variety of situations (e.g. Zaragoza & Lane, 1994; Crombag, Wagenaar, & Van Koppen, 1996).

A second type of study employs the strategy of repeated interviews about non-events to create pseudomemories in subjects. Again, it can be shown that children as well as adults tend to develop pseudomemories in response to such interviews (e.g. Ceci, Crotteau Huffman, Smith, & Loftus, 1994; Hyman, Husband, & Billings, 1995).

Quite another approach was proposed by Barclay and Wellman (1986; see also Barclay, 1986; Barclay & Decooke, 1988). These researchers had subjects record autobiographical events over an extended period of time. Following this, subjects underwent recognition tests in which they had to differentiate between original records and foils fabricated by the authors. In general, the results of this study indicate that people are good at recognizing original autobiographical fragments, but relatively poor at rejecting pseudomemories (i.e. foils). That is, a relatively high percentage of foils was accepted as authentic memories and the more so when the time interval between autobiographical notes and recognition tests became longer.

So far, authors concerned with memory reliability and psychotherapy have largely ignored the study by Barclay and Wellman (1986). Yet, their study might be important because it suggests that even in the absence of post-event misinformation and repeated interviews about non-events, autobiographical recollections may become surprisingly inaccurate. The primary aim of the present study was to examine whether the findings obtained by Barclay and Wellman (1986) represent reliable phenomena. Thus, a replication study was carried out: undergraduate students kept autobiographical notes over a one-week period and, after a four-month interval, underwent a recognition task. A second issue addressed by the current study was the association between psychopathological symptom levels and accuracy of autobiographical memory. There are good reasons to believe that memory illusions may occur in the context of psychotherapy (e.g. Loftus, 1993). Therefore, the question arises whether people with relatively high levels of depression, anxiety or dissociation are more inclined to accept foils as accurate memories. Given the small sample size, this issue could only be addressed in a preliminary fashion.

## **Method**

### *Subjects*

Subjects were 10 female undergraduates in Health Sciences at the University of Limburg. Their mean age was a 25.8 years (range: 22–32 years). Subjects were paid for their participation.

### *Procedure*

Subjects were asked to provide brief, but complete descriptions of three self-selected, outstanding events every day, for seven consecutive days. In describing the events, subjects used a fixed format. That is, they briefly specified "context" (i.e. place), "event" (i.e. what happened), and "affective evaluation" (i.e. emotional reaction).

After four months, subjects were invited to participate in a wordlist study. Subjects were not informed about the pertinent recognition questionnaire (see below). Following the wordlist experiment, subjects underwent a recognition test. More specifically, each subject completed a questionnaire consisting of 30 items: 10 items were literal transcriptions of authentic reports provided by that subject and 20 items were foils. There were three types of foils: 5 foil items pertained to original records of the subject in which only the context was radically changed (context-foils; e.g. "driving to university" became "driving to my parents"); 5 items involved original records of the subject in which only the evaluation was radically changed (evaluation-foils; e.g. "I was afraid" became "I was angry"); and 10 items were authentic reports of another subject (other-foils). Foils manufactured by the experimenters deviated strongly from the original records, but had sufficient *prima facie* plausibility. That is, totally implausible modifications (e.g. "driving to the Himalaya") as well as vague and possibly correct paraphrases (e.g. "driving to school") of the original records were avoided. Using a yes/no format, subjects indicated whether or not items were exact duplicates of their original memories. In addition, they rated on a 100 mm Visual Analogue Scale (VAS) how confident they were about each answer (0 = not confident at all; 100 = absolutely confident). The recognition questionnaire took about 15 minutes. Following this, subjects completed three questionnaires: the Beck Depression Questionnaire (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), the Fear Questionnaire (FQ; Marks & Mathews, 1979), and the Dissociative Experience Scale (DES; Bernstein & Putnam, 1986). The BDI is a self-rating questionnaire consisting of 21 items. It measures the extent to which behavioural manifestations of depression are present. The FQ is a 15-item questionnaire which asks for phobic symptoms. The higher the total FQ score, the higher the level of phobic symptomatology. The DES is a 28-item instrument that presumably measures dissociation. The higher the DES score, the stronger the tendency to dissociate.

## Results

Table 1 presents mean proportions of authentic memories that were accurately identified (i.e. hits) and mean proportions of foils that were classified as authentic memories (i.e. false alarms). The right column of Table 1 shows the mean VAS-confidence ratings for hits and false alarms. As can be seen, 78% of the authentic memories was accurately classified (i.e. 78 out of 100 original notes). A nonparametric Friedman test showed that the proportions of false alarms for the three types of foils differed significantly,  $X(2) = 8.5$ ,  $p < .02$ . Separate Friedman tests indicated that context-foils gave rise to more false alarms than the other-foils,  $X(1) = 8.1$ ,  $p < .01$ , whereas evaluation-foils and other-foils did not differ,  $X(1) = 1.6$ ,  $p = .21$ . Also, context-foils and evaluation-foils did not differ with regard to the proportion of false alarms,  $X(1) = 1.5$ ,  $p = .25$ . For the sake of completeness, mean discrimination and bias index ( $Pr$  and  $Br$ , respectively) were calculated, using the formulas proposed by Snodgrass and Corwin (1988).  $Pr$  was .78 and  $Br$  was .55. The latter value comes close to .5 and, therefore, indicates a relatively neutral response bias.

Table 2 provides some examples of original memories that were not recognized (i.e. misses) and foils that were accepted as authentic memories.

TABLE 1. Proportions of hits (originals) and false alarms (foils) and confidence ratings (standard deviations are given between parentheses)

Items	Hits/false alarms (0-1.0)	Confidence (0-100)
Originals	.78	92 (2.4)
Context-foils	.34	83 (5.6)
Evaluation-foils	.20	92 (2.8)
Other-foils	.03	80 (2.0)

TABLE 2. Examples of misses and false alarms

### Misses

At a student meeting. There was discussion about a suicide attempt. They kept on talking about this issue. I was fascinated.

At the travel shop. I was there with a friend to arrange a flight to Spain. I was completely happy.

### False alarms (context changed)

At my friend's apartment (original: at home). My boy friend called me to say that he was ill and could not come along to the party. I was disappointed.

On my way downtown with a good friend (original: I went with my friend to the

museum, outside the city). She told me a lot of personal things, about her illness. I appreciated that very much.

**False alarms (evaluation changed)**

I was at a party. Some guys drank too much. They had a row. I was amused. (original: I found it childish).

At the theatre. We were playing a scene. It failed completely and the other scenes were also rather bad. I felt indifferent about it. (original: I was very sad and almost started to cry).

**False alarms (others)**

I went to the city and I crossed a busy road. I nearly was hit by a car. I was very alarmed.

At home. I received a nice postcard. I was moved.

---

Confidence ratings were evaluated with a within-subject Analysis of Variance (ANOVA). The confidence ratings for the other-foils were excluded from this analysis, because they were based on only a small number of observations (3 out of 100 other-foils). The ANOVA indicated that confidence ratings for hits and false alarms did not differ significantly,  $F(2, 10) = 2.1, p = .19$ .

Spearman rank correlations were computed between total number of false alarms, on the one hand and levels of depression, phobic anxiety, and dissociation, on the other hand. Except for the correlation between BDI score and false alarms (Spearman  $r = .56, p < .05$ ), no correlation attained significance (Spearman  $r$ 's  $< .42, p$ 's  $> .11$ ).

## Discussion

The main results of the current study can be summarized as follows. First, the present study replicates previous findings reported by Barclay and Wellman (1986). That is, after an interval of several months, both acceptance of original memories and rejection of foils were not perfect. Interestingly, false alarms were not randomly distributed over the three types of foils: other-foils elicited fewer false alarms than did context-foils. This is broadly in line with the pattern found by Barclay and Wellman (1986; see also, Barclay, 1986, Table 6.2). Yet, it is worthy of note that the false alarm rates reported by these authors are on the whole higher than those obtained in the current study (e.g. at a 4–6 month delay, Barclay and Wellman found false alarm proportions for evaluation-foils and other-foils of .47 and .30, respectively). Possibly, this is due to the fact that in the current experiment, foils were manufactured in such a way that they strongly deviated from the

original, whereas such a procedure was apparently less prominent in the Barclay and Wellman study (see, for an extensive discussion of this issue, Barclay & DeCooke, 1988).

Second, confidence ratings for false alarms did not differ from confidence ratings for hits. Again, this finding confirms the results of Barclay and Wellman (1986), who also found that false alarms are accompanied by high confidence scores.

Third, the current study found that depressive symptoms, but not levels of dissociation or phobic symptoms, are associated with a higher frequency of false alarms. While these results have a preliminary status, they are in accordance with studies demonstrating autobiographical deficits in depression. Germane to this issue is the work by Williams (1992) who showed that depressive individuals are characterized by global and poorly detailed autobiographical memories. It may well be the case that false alarms can only occur against a background of vague memories. The connection between depression, global memories, and false alarms warrants systematic investigation, especially because memory illusions about significant events may occur in people suffering from psychopathology (e.g. Loftus, 1993; Lipinsky & Pope, 1994).

By and large, the present findings are consistent with Barclay and Wellman's (1986) hypothesis that autobiographical memory is guided by schematized reconstruction. Such reconstruction is not necessarily correct, which is demonstrated by misses and false alarms. To the extent that false alarms on a recognition test can be interpreted in terms of pseudomemories, the present findings together with those of Barclay and Wellman (1986) suggest that pseudomemories are not critically dependent on repeated interviews or post-event misinformation. Pseudomemories may have an internal source, i.e. pseudomemories may originate from the schematized reconstruction that underlies autobiographical memory. In Barclay and Wellman's (1986, p. 101) words: "What one remembers then is, at least in part, what could have happened or should have happened in one's life". In a sense, the false alarms obtained with the Barclay and Wellman procedure parallel the recent findings of Roediger and McDermott (1995). In a typical wordlist experiment, these authors gave subjects series of words that were semantically related to each other (e.g. bed, rest, etc). On a post-recognition test, these target words were presented along with foils that were associated with the target words (e.g. sleep), but never appeared in the original series. Under these conditions, subjects tend to react with a relatively high frequency of false alarms to the foils. Accordingly, the authors concluded that their results "reveal a powerful illusion of memory: people remember events that never happened" (Roediger & McDermott, 1995, p. 803).

Some comments on the methodological and conceptual limitations of the present study are in order. In laboratory studies such as those of Roediger and McDermott (1995), experimenters have complete control over foils, i.e. they can guarantee that foils are, indeed, foils. Clearly, such a degree of certainty can not be reached with the paradigm proposed by Barclay and Wellman (see, for a critical discussion, Brewer, 1988; McCauley, 1988). Thus, even though subjects are instructed to evaluate whether items represent exact duplicates of their own records, there is a remote possibility that foils capture an event that subjects actually experienced. In that case, it would be misleading to talk about "false alarms" and "pseudomemories". One way to settle this issue is to interview subjects about their acceptance of foil items. A second point concerns the content of the autobiographical notes. In the current experiment, subjects were asked to record outstanding events. Likewise, in the Barclay and Wellman (1986) study, subjects were instructed to record "memorable" events. Despite these instructions, it remains possible that subjects recorded routine events (Barclay & DeCooke, 1988; McCauley, 1988). If true, misses and false alarms would be a function of the low significance of the recorded events. In that case, they would have no relevance to the type of memories that are retrieved during psychotherapy, because these memories usually pertain to crucial experiences. Inspection of the autobiographical records in the current study suggests that they often referred to outstanding events and only rarely to trivial situations (see e.g. Table 2). Nevertheless, this conceptual point requires further study in order to establish to what extent memory illusions about outstanding events are relevant to clinical practice in general. Obviously, memory illusions of the sort found in the current study can not be generalized to the special case of traumatic memories recovered during psychotherapy (see, for a discussion, Freyd & Gleaves, 1996; Roediger & McDermott, 1996). Pertinent to this issue are studies that suggest that traumatic memories may have a different representation than verbal accessible memories (e.g. Brewin, Dalgleish, & Joseph, 1996).

Despite these limitations, the Barclay and Wellman (1986) approach appears to be a promising technique for studying individual difference factors in the development of memory illusions. More detailed knowledge about these factors is clearly important for clinical practice.

## References

- BARCLAY, C.R. (1986). Schematization of autobiographical memory. In D.C. Rubin (Ed.), *Autobiographical memory* (pp. 82-99). Cambridge: Cambridge University Press.
- BARCLAY, C.R. & DeCooke, P.A. (1988). Ordinary everyday memories: Some of



- the things of which selves are made. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 91-125). Cambridge: Cambridge University Press.
- BARCLAY, C.R. & WELLMAN, H.M. (1986). Accuracies and inaccuracies in autobiographical memories. *Journal of Memory and Language*, 25, 93-103.
- BECK, A.T., WARD, C.H., MENDELSON, M., MOCK, J.E., & ERBAUGH, J.K. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- BERNSTEIN, E.M. & PUTNAM, F.W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease*, 174, 727-735.
- BREWER, W.F. (1988). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 21-90). Cambridge: Cambridge University Press.
- BREWIN, C.R. (1996). Scientific status of recovered memories. *British Journal of Psychiatry*, 169, 131-134.
- BREWIN, C.R., DALGLEISH, T., & JOSEPH, S. (1996). A dual representation of post-traumatic stress disorder. *Psychological Review*, 103, in press.
- CECI, S.J., CROTTEAU HUFFMAN, M.L., SMITH, E., & LOFTUS, E.F. (1994). Repeatedly thinking about a non-event: Source misattributions among preschoolers. *Consciousness and Cognition*, 3, 388-407.
- CROMBAG, H.F.M., WAGENAAR, W.A., & VAN KOPPEN, P.J. (1996). Crashing memories and the problem of source monitoring. *Applied Cognitive Psychology*, 10, 95-104.
- FREYD, J.J. & GLEAVES, D.H. (1996). Remembering words not presented in lists: Relevance to the current recovered/false memory controversy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 811-813.
- HYMAN, I.E., HUSBAND, T.H., & BILLINGS, F.J. (1995). False memories of childhood experiences. *Applied Cognitive Psychology*, 9, 181-197.
- LINDSAY, D.S. & READ, J.D. (1994). Psychotherapy and memories of childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology*, 8, 281-338.
- LIPINSKY, J.F. & POPE, H.G. (1994). Do flashbacks represent obsessional imagery? *Comprehensive Psychiatry*, 35, 245-247.
- LOFTUS, E.F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
- LOFTUS, E.F. (1993). The reality of repressed memories. *American Psychologist*, 48, 518-537.
- MARKS, I.M. & MATHEWS, A.M. (1979). Brief standard self-rating scale for phobias. *Behaviour Research and Therapy*, 17, 263-267.
- MCCAULEY, R.N. (1988). Walking in our own footsteps: Autobiographical memory and reconstruction. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 126-144). Cambridge: Cambridge University Press.
- PARIS, J. (1995). Memories of abuse in borderline patients: True or false? *Harvard Review of Psychiatry*, 3, 10-17.

- ROEDIGER, H.L. & McDERMOTT, K.B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 803-814.
- ROEDIGER, H.L. & McDERMOTT, K.B. (1996). False perceptions of false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 814-816.
- SNODGRASS, J.G. & CORWIN, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34-50.
- TRIERWEILER, S.J. & DONOVAN, C.M. (1994). Exploring the ecological foundations of memory in psychotherapy: Interpersonal affordance, perception, and recollection in real time. *Clinical Psychology Review*, 14, 301-326.
- WILLIAMS, J.M.G. (1992). Autobiographical memory and emotional disorders. In S.A. Christianson (Ed.), *Handbook of emotion and memory: Research and theory* (pp. 451-477). Hillsdale, NJ: Erlbaum.
- ZARAGOZA, M.S. & LANE, S.M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1-12.